Received: September 27, 2023. Revised: November 17, 2023.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Enhancing Hybrid System Based on Reinforcement Learning

Zeena Mustafa[1]* **Ekhlas Kadhum Hamza[1]**

[1]*Department of Control and Systems Engineering, Technology University, Baghdad, Iraq*
* Corresponding author's Email: Ekhlas.K.Hamza@uotechnology.edu.iq

Abstract: Light fidelity (Li-Fi) can be defined as a type of wireless technology that sends data via light waves through LED light bulbs. Light fidelity (Li-Fi) offers high-speed data transmission capabilities and a large unlicensed bandwidth, making it a promising technology for the future. However, factors including interference, wall reflection, and blocking may cause the quality of a light fidelity (Li-Fi) channel to vary from one part of a room to another. Another type of wireless communication technology that offers broad coverage and slow transmission rates is wireless fidelity (Wi-Fi). Since the electromagnetic spectrums regarding such two technologies do not overlap, there is a possibility for building a hybrid Wi-Fi and Li-Fi network that provides seamless and high throughput global communication. One wireless fidelity access point (Wi-Fi AP) and four light fidelity access points (Li-Fi Aps) make up the downlink hybrid system we discuss in this work. Finding an access point (AP) assignment method that will increase long-term system throughput at the same time as still guaranteeing users' pleasure and fairness is challenging. Thus, the authors suggest using a reinforcement learning (RL) algorithm. The algorithm aims to balance the load of multiple access points (APs) by considering both the LiFi and Wi-Fi channels. The results obtained using MATLAB code for a hybrid system based on reinforcement learning (RL) and a standard solution set size (SSS) access technology called TDMA. The system was evaluated in terms of throughput and user satisfaction for 5 users. According to the results, the RL-based hybrid system achieved a throughput of up to 210 Mbps and an SSS of 180 Mbps. Additionally, the user satisfaction was reported to be 100%.

Keywords: Time division multiplexing (TDM), Complementary cumulative distribution function (CCDF), Random waypoint (RWP), Light fidelity (Li-Fi), Wireless fidelity (Wi-Fi).

1. Introduction

Machine learning (ML) known as reinforcement learning (RL) is concerned with teaching autonomous agents to respond in response to incentives and punishments. Positive outcomes from using RL in areas such as network optimization, traffic engineering, and resource allocation have been observed recently. Hybrid Li-Fi and Wi-Fi networks, which are quickly gaining popularity as significant technology, combine the advantages of the two to enable quick wireless data transfer. While Wi-Fi offers moderate data rates and near-universal coverage, Li-Fi employs LED light bulbs to transmit data using light waves. Nevertheless, interference, reflected signals from walls, and obstructions can all cause the Li-Fi channel's quality to fluctuate. When operating a hybrid network, it is essential to distribute traffic evenly across all of the APs so that you can maintain a constant and high data rate for your users. As a result of their ability to adapt to their surroundings, RL-based load-balancing algorithms have garnered a lot of interest for use in this scenario.

We provide an RL-based load balancing technique about a down-link hybrid system with one Wi-Fi AP and four Li-Fi APs in this paper. The objective is to increase overall system throughput in a manner that is fair to and well-liked by all users. The suggested approach considers both Wi-Fi and Li-Fi channels, considering each channel's properties and the load on each one of the access points for making intelligent decisions regarding AP assignments. Signal-strength method, iterative optimization, and exhaustive search are just a few of the cutting-edge benchmark methods we compare the suggested RL-based algorithm against. As an

example, we test the proposed method using the random waypoint model in non-uniform as well as uniform user distribution. According to the results, the proposed RL-based algorithm is frequently better than benchmark methods, making it a great option for load balancing in hybrid Wi-Fi Li-Fi networks. We believe that the proposed approach will work well for load balancing in scenarios, in which the Li-Fi channel's quality varies for a number of different causes.

Due to the rising demand for wireless data, Cisco projects that by the year of 2022, monthly global mobile data traffic will reach a staggering 396 Exabytes (396billion GB). Seven times as many were found in 2018 as in 2017 [1]. Visible light communication (VLC) for interior spaces is being investigated by academic and commercial researchers to help address the rising demand for wireless data. VLC has various benefits over conventional RF communication, including a wider range of usable frequencies, the ability to reuse existing infrastructure, and a higher level of security. In addition to being less disruptive to other electronic equipment, VLC also does not pose any health risks [2, 3]. Additionally, VLC is considered a safer alternative as it does not pose any health hazards and does not interfere with other electro-magnetic devices [4].VLC relies on both direct detection (DD) and intensity modulation (IM) to function. Modulating the light signal's intensity allows for the transmission of data, and at the receiving end, fluctuations in the received intensity are detected and encoded. LEDs are used by VLC to transport data, and a photodetector (PD) is used to convert received optical signal back to electrical signal. Overall, VLC shows promise as a technology to fulfil the future data needs because to its capacity to deliver high-speed and secure communication for indoor environments. LiFi makes use of VLC's physical layer as a means of communication. It's made such that high-velocity, fully-networked, two-way wireless communication is possible. LiFi, in essence, is a wireless networking capability [5, 6]. provided by extending point-to-point VLC. Several researchers have found that well-designed hybrid LiFi and Wi-Fi network can provide increased data speeds, better outage efficiency, and happier users [7, 8]. Finding the optimal AP homework assignment has been the subject of numerous studies. Strategies for hybrid LiFi and Wi-Fi networks, which are outlined below.

The issue, is the challenges that numerous. Here is a brief description of the potential drawbacks associated with each technique: Fuzzy logic may struggle to adapt to dynamic and changing environments as it relies on predefined fuzzy rule sets,difficulty in rule design: Constructing accurate and comprehensive fuzzy rule sets can be challenging and time-consuming, lack of optimization, and fuzzy logic may not inherently optimize system performance or learning. RL method: Exploration-exploitation trade-off: RL methods often face the challenge of balancing exploration (searching for optimal actions) and exploitation (taking advantage of known good actions) to maximize performance, and High computational complexity: RL algorithms can be computationally demanding and may require extensive training time and computational resources.By highlighting these drawbacks, can emphasize the unique features and potential advantages of proposed approach for improving performance in hybrid systems based on reinforcement learning.

1.1 Related work

Studies have shown that reinforcement learning (RL) algorithms have great potential in improving the performance of wireless networks, particularly in terms of optimizing resource allocation and load balancing. Yet, little research has been done on particularly using RL algorithms in hybrid LiFi as well as Wi-Fi networks. One research suggested an access point (AP) assignment optimization genetic algorithm as a load balancing approach for hybrid LiFi and WiFi network. The algorithm aims to balance the load across different APs while maximizing the network capacity and ensuring user fairness. With regard to a hybrid LiFi and Wi-Fi network, a different study suggested a dynamic programming solution for load balancing. The method models the network and chooses the best AP assignment plan using Markov decision process (MDP). There is related work on using reinforcement learning (RL) algorithms for optimizing resource allocation and load balancing in wireless networks, but there is limited work specifically on applying RL to hybrid LiFi and WiFi networks. There are work that apply RL algorithms in cellular and Wi-Fi networks, among other kinds of wireless networks. These studies demonstrate the potential of RL in improving network performance, user satisfaction, and resource utilization. Particle swarm optimization (PSO) was used by the authors of. to suggest an approach for regulating mobility and distributing resources in indoor VLC networks [9, 10]. Those methods' significant computational complexity results from the fact that each time-step an optimization issue should be addressed to implement them.

Other research has also used FL to allocate APs

to participants. FL-based dynamic load balancing technique was presented by [11] to reduce impacts of handover, which would lessen the influence of handoffs. In the case when determining whether or not a handover is required, this technique takes the user's selected data rate and transfer speed into account. Users moving quickly or experiencing temporary shadowing effects will have better APs assigned to them based on the speed data provided by FL. The plan also stops handoffs from happening in a ping-pong pattern. Another research suggested a two-step AP selection technique in to address this issue. With regard to the first stage, FL is utilized to allocate users to Wi-Fi access points, and in 2nd stage, the rest of the users are allocated to Li-Fi network. This approach required substantially less work to attain the same throughput as the optimization-based approach.

The viability of applying ML to address the problem of AP assignment in Li-Fi/Wi-Fi hybrid networks has been investigated in various research. As an illustration, the authors of introduced the concepts of responsive association (RA), which classifies users depending on their present geo-locations as well as queue backlog states, and anticipatory association (AA), which takes into account their time-varying geo-locations and shifting queue backlog states. The authors shared their findings regarding the compromise between latency and throughput. Reinforcement learning with knowledge transfer was proposed as the basis for a network selection algorithm by the authors of [12, 13]. Context information and the stationary distribution law regarding network load are utilized for aiding in the construction of algorithm to satisfy the traffic's asymmetric uplink and downlink performance needs. Reward function, which is a function related to the instantaneous downlink and uplink throughput, rather than other QoS factors like user fairness and satisfaction, was the only one examined in this paper. Multi-armed bandit algorithm is used by the authors of [14] to present AP selection approaches. Decision probability distributions are updated with use of "exponential weight values for explorations and exploitations of the "algorithm and "exponentially weighted algorithm with the linear programming" algorithm. The approach that has been employed in the suggested study differs from earlier works since TRPO is utilized in place of Q-Learning with knowledge transfer in the present study. Along with evaluating user satisfaction and fairness, our study also aims to improve the parameter. The suggested technique varies from the existing work in that it uses multi-agent learning rather than single-agent luminance descent-based learning.

Li-Fi technology has been refined and advanced through numerous scholarly and global research projects. The efficiency, availability, security, and safety of light fidelity turn today's telecommunication into tomorrow's visible light communication as new wireless communication technology advances to use LED.

The current radio frequency (RF) networks are under threat from the growing number of mobile devices. To reduce the spatial variation in data rate, a hybrid radio frequency/visible light communications (HLRN) network (HLRN) is suggested, which provides a higher system throughput than independent radio frequency (RF) or visible light communications (VLC) networks. The primary issue with hybrid networks is load balancing, which degrades network speed. As a result, in order to address this issue, the load balancing (LB) schemes in HLRNs are examined with an emphasis on the users' AP assignment. In that case, it is discovered that effective spectrum sensing—which uses five technologies split into three stages—is necessary to prevent the incorrect holes in the band from being detected. These technologies are received signal strength (RSS), particle swarm optimization (PSO), and deep learning techniques. Convolutional, feed forward neural network (FFNN) [28].

We developed a RL technique for load balancing in a network which integrates Wi-Fi and Li-Fi as a result of the research's findings. Our method succeeds in a number of established categories, including average network performance, user satisfaction, fairness, and outage resilience. To improve user happiness and network equity, we have developed a reward mechanism for the RL algorithm that maximizes long-term average network throughput. We compare our results to the most effective iterative optimization, best available signal strength strategy (SSS), and exhaustive search. Results are shown with regard to typical network speed, computer complexity, user satisfaction, fairness, and how frequently capacity outages occur. We have also considered two different user behavior models: hotspot random waypoint (HRWP) and random waypoint (RWP) for demonstrating the adaptability of our proposed RL technique.

Our main contributions are summarized in the following way:

- For the hybrid LiFi and WiFi networks, we presented RL-based technique of dynamic load balancing.
- We have designed a reward function that aims to improve both long-term network throughput and users' satisfaction and

fairness.

- We have contrasted our suggested approach with leading-edge approaches and reported the findings in terms of several performance metrics.
- To show how flexible our proposed approach is, we've taken into account two scenarios, each with its own model of user behaviour.

The remainder of this essay has been organized in the following manner: We analyze system model in sections 2 and 3, we present our suggested RL-based load balancing approaches. Part 4 presents the evaluation and discussion of the performance, and section 5 wraps up the study.

2. System model

The hybrid LiFi and WiFi network that supports numerous users for indoor communication is the main topic of this article. Four LiFi APs and one Wi-Fi AP, which are restricted to smaller attocell zones, make up the network. Nu and NAP, which stand for total number of the users and APs, respectively, serve as the system's representations. Consider a typical room with the dimensions 5 x 5 x 3 m^3, where the WiFi AP is placed in the middle in order to cover the entire area. A central unit (CU) connected to LiFi as well as Wi-Fi APs makes the load balancing decisions. CU has access to the accurate feedback data required to decide how to distribute the load. In the overlapping attocell areas, optical interference may occur due to the reuse of the same modulation bandwidth by all LiFi APs.

The interference has been treated as background noise in this study. Users are evenly spread over the area and move around according to random way-point and hotspot random way-point models. Stream high-definition videos online, clients require greater data rates that have been simulated as poisson process with a parameter that has been set to the value of 50Mbps. In the presented study, multiple user connections to a single AP are made possible by the use of time-division multiple access (TDMA), and users are only able to connect to one AP (either the Wi-Fi AP or the LiFi AP) at a time. Resources are distributed via a round-robin approach.

A. LiFi channel model

The behavior of the light signal conveyed between photodetector receiver and LED light source is described mathematically by the LiFi channel model. The channel model accounts for a number of variables, including the separation between the receiver and transmitter, the presence of obstructions

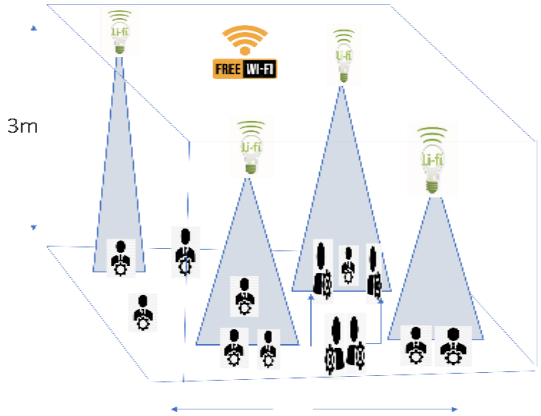

Figure. 1 Diagram of hybrid LiFi Wi-Fi network

In light's path, the angle at which it is incident, and the characteristics of light source and receiver, such as their field of vision and spectrum sensitivity. The channel model is frequently represented mathematically as an equation that connects the received optical power to transmitted optical power and other system variables like the receiver field of view, the transmission distance, and the medium's extinction coefficient (i.e., the attenuation of the light signal caused by scattering and absorption by the medium). There are several types of channel models used in LiFi systems, such as the Lambertian channel model, the log-normal channel model, and the Ricean channel model, each with its own assumptions and parameters. The choice of channel model depends on the specific application and the characteristics of the LiFi system.

The optical channel of LiFi is composed of two parts, which are non-line-of-sight (NLOS) component and the line-of-sight (LOS) component. The channel gain of a LOS component [8] can be expressed as follows:

$$HLOS = \frac{(m+1)A_{PD}}{2\pi d^2} cos(\varphi) g_f \, g_c(\psi) \, cos(\psi) \quad (1)$$

It is possible to repeat the formula "in which A_{PD} represents PD's physical area, m denotes Lambertian order related to the transmitter, d represents a distance between the user and LiFi AP, gf represents optical filter gain, and gc represents optical concentrator" as:

$$f(x) = \begin{cases} \frac{n^2}{sin2(\psi)}, 0 \leq \psi \leq \Psi \\ 0, \psi > \Psi \end{cases} \quad (2)$$

The channel gain for the NLOS component can be defined as [10, 15], where Ψ represents the semi-angle of the field of view (FOV) of PD, and n denotes the reflective index.

Table 1. Channel parameters of LiFi

Parameters of the Channel	Symbols	Values
Responsivity	R_{PD}	0.53 A/W
Reflection coefficient	ρ	0.8
Transmit optical power per LiFi AP	P_{opt}	3 Watt
Band-width/LiFi AP	B_{LiFi}	40MHz
PSD of the Li-Fi noise	N_{LiFi}	$10^{-21}A^2$
Height difference between user and AP	h	2m
PD Area	A_{PD}	1 cm^2
Gain of optical filter	g_f	1
Half intensity radiation angle	$\theta_{1/2}$	60°
PD Field of View (FOV)	Ψ	60°

$$H_{NLOS} = \frac{\rho A_{PD} e^{j2\pi f \Delta T}}{A_{room}(1-\rho)(1+j\frac{f}{fc})} \quad (3)$$

In this work, we have used a channel model that takes into account the delay (ΔT) between diffused signals and line-of-sight (LOS), the room area (A_{room}), and cut-off frequency (fc), along with the reflectivity (ρ) of the walls. Although our model can incorporate non-constant reflectivity from various surfaces, for simplicity, we assumed a constant reflectivity for the reflecting surfaces in this study. However, it should be noted that this simplification does not alter the conclusions of our research [16].

The optical channel can be divided into two components, namely HLOS and HNLOS, which together make up the complete optical channel H Li-Fi. The parameters of simulation that have been utilized for the Li=Fi channel have been provided in Table 1. SNR of the user μ, who is connected to the LiFi access point α, may be denoted as SNRμ, α and is given by the following expression.

$$SNR_{\mu,\alpha} = \frac{\left(H_{LiFi(\mu,\alpha)} P_{opt} R\right)^2}{N_{LiFi} B_{LiFi}} \quad (4)$$

SNR, for a user who is connected to a LiFi AP depends on several factors, including the AP's and the user's channel gain, the user's PD responsivity R, LiFi noise power spectral density (PSD) H_{LiFi}, and optical power transmitted, the AP's band-width, or B_{LiFi}. Any signal received from a different LiFi AP will be perceived as interference since each LiFi AP reuses the same frequency. Thus, the notation SINR indicates the user's signal-to-interference-noise ratio (SINR) while connected to a LiFi AP and is equal to.

Table 2. Channel parameters of Wi-Fi

Channel Parameters	Symbols	Values
Transmit Power	$P_{Wi\text{-}Fi}$	20 dBm
Shadowing loss	X_{SF}	3dB
Bandwidth per Wi-Fi AP	$B_{Wi\text{-}Fi}$	20MHz
PSD of noise	$N_{Wi\text{-}Fi}$	-174dBm/Hz
Breakpoint distance	d_{BP}	5cm
Central carrier frequency	fc	2.4 GHz

The SINR for user μ who is connected to Li-FiAP α is hence referred to as SINRμ and is calculated as follows:

$$SNR_{\mu,\alpha} = \frac{\left(H_{LiFi(\mu,\alpha)} P_{opt} R\right)^2}{N_{LiFi} B_{LiFi} + P_{\beta \in AP}\left(H_{LiFi(\mu,\alpha)} P_{opt} R\right)^2} \quad (5)$$

H_{LiFi} (μ, β) represents the gain of the channel between the user μ and interfering LiFi APs β. The capacity lower bound is utilized for calculating the possible data rates between the Li-Fi AP α and the user μ and presents the next results:

$$r_{\mu,\alpha} = \frac{B}{2} log_2\left(1 + \left(\frac{6}{\pi e}\right) SINR_{\mu,\alpha}\right) \quad (6)$$

B. Wi-Fi channel model

Here WiFi channel can be given as:

$$G_{\mu,\alpha}(f) = \sqrt{10^{-\frac{L(d)}{10}}}\, hr \quad (7)$$

The equation, where f stands for the carrier frequency, represents both large-scale fading loss L (d) as well as small-scale fading gain hr (in dB).

With a mean value of the power of 2.46 dB, small-scale fading gain hr follows independent identical Rayleigh distributions, while L (d) [17] can be determined by:

$$L(d) = \begin{cases} L_{FS}(d) + X_{SF}, d < d_{BP} \\ L_{FS}(d_{BP}) + 35log\left(\frac{d}{d_{BP}}\right) + X_{SF}, d \geq d_{BP} \end{cases} \quad (8)$$

LFS stands for the free space loss, whereas d stands for the separation distance. The shadowing loss is represented by XSF, while the breakpoint distance is marked by dBP. LFS (d) = 20log10 (d) + 20log10 (f) 147.5dB, in which f is the frequency,

yields the free space loss. Wi-Fi channel parameters that have been used in the simulation have been listed in Table 2. A user (μ) connecting to a Wi-Fi access point's SNR could be expressed as follows:

$$SNR_{\mu,\alpha}(f) = \frac{|G_{\mu,\alpha}|^2 (f) P_T}{N_{WiFi} B_{WiFi}}, \quad (9)$$

The transmitted power is denoted by P_T, the noise PSD in the WiFi link is represented by $N_{Wi\text{-}Fi}$, and the bandwidth of the Wi-Fi access point is denoted by $B_{Wi\text{-}Fi}$. The gain G (μ, α) (f) is determined with the use of Eq. (7). There won't be any interference because there is only one WiFi access point in the system, and Eq. (9) states that SNIR for the Wi-Fi access point will be equal to SNR. Using Shannon's capacity formula, the attainable data rate between the Wi-Fi access point and the user is calculated, giving the next statement:

$$r_{\mu,\alpha} = B \log_2(1 + SNR_{\mu,\alpha}) \quad (10)$$

C. Random waypoint mobility model

The majority of mobility research employs a a random waypoint (RWP) paradigm. In this scenario, the user chooses a location at random and proceeds towards it at a steady rate. When the final objective has been reached, the user chooses a new location and proceeds in that direction. Nevertheless, it's also important to think about the existence of hotspots within a room, which are areas where people concentrate with a high probability. For instance, people may cluster near LEDs to overall lighting quality.

To generate attraction nodes, the authors of [18], present a variant of the random waypoint model in which the distribution of final stops can be altered. $X_{a,\ min}$, $y_{a,\ max,}$, and $y_{a,\ min}$ determine the region of attraction point A_a, whereas the intensity is expressed by.

The given distribution for destination s points (x_d, y_d) is as follows:

$$f(x_d, y_d) = \frac{1}{A_{room} + (\xi - 1)A_a} \times [(u(x_d + x_m) - u(x_d - x_m))(u(y_d + y_m) - u(y_d - y_m)) + (\xi - 1)(u(x_d - x_{a\,max}) - u(x_d - x_{a\,min}))(u(y_d - y_{a\,max}) - u(y_d - y_{a\,min})) \quad (11)$$

The attraction point area is denoted by A_a and is defined by the conditions $x_{a,\ min} \le x \le y_{a,\ max}$ and $y_{a,\ min} \le y \le y_{a,\ max}$, while the intensity is represented by ξ.

In our experiment, we used a value of ξ = 200 and an area of $A_a = 0.25\ m^2$ to create four hotspots centered on four LEDs placed across the room. We've also built in a pause allowing users to rest at a hotspot

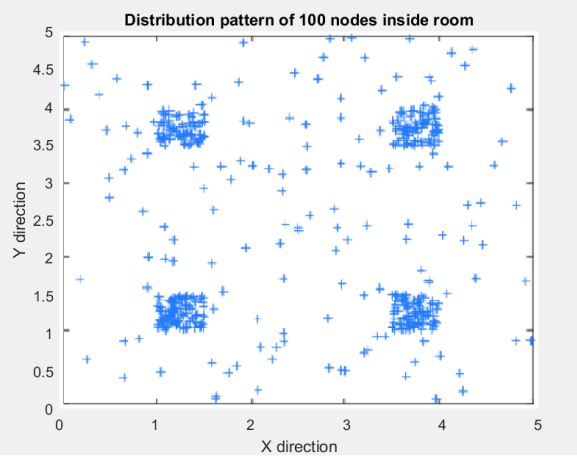

Figure. 2 Pattern of distribution of 100 nodes inside a room for the modified random way-point model

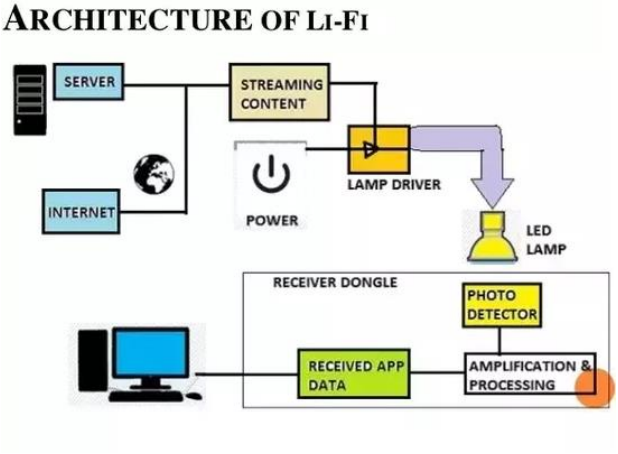

Figure. 3 Hybrid LiFi WiFi

for a certain amount of time before proceeding to a new location. In this paper, we'll refer to this specific mobility model as the HRWP. Based on the foregoing requirements for the HRWP model, the placement of 100 nodes in the room is shown in Fig. 2.

3. Suggested load balancing method

A. Reinforcement-learning (RL) approach

An agent interacts with the environment continuously to observe its state and carry out actions as a response to such observations. This significant ML method is shown in Fig. 3 as RL. An RL agent can map states to the distribution of probabilities over the actions for the maximization of the cumulative reward. In the lack of information on MDP, one can consider RL to be a stochastically ideal solution to the MDP. According to the standard formulation of MDP, an agent is in state st ∈ S at time step t ≥ 0, performs an action at ∈ A, receives instantaneous reward at rt = r (s_t, a_t) R, and then switches to state st+1P (.|s_t, a_t) S.

The formula for the policy is π: S → P (A), where P (A) stands for the the set of the distributions over

action space A [21]. According to policy π, the reduced cumulative award is:

$$\eta(\pi) = E_{s0,a0,\ldots}[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t)], s_o \sim \rho_o(s_o), \quad a_t \sim \pi(a_t|s_t), s_{t+1} \sim P(s_{t+1}|s_t, a_t) \quad (12)$$

The emphasis on immediate benefits is determined by the discount factor γ ∈ (0, 1), with lower values of factor giving greater weight to immediate rewards. RL objective is identifying optimum policy, π*, which maximizes η (π), where:

$$\pi^* = arg\_max\,(\eta(\pi)) \quad (13)$$

We used infinite-horizon discounted MDP, defined as (S, A, P, r, ρ0, γ), to represent the given hybrid LiFi WiFi network AP assignment problem.

• To represent the assignment of Aps in a hybrid LiFi/Wi-Fi network, infinite-horizon discounted MDP with the parameters (S, A, P, r, 0) was used.

S represents a set of the continuous states including the SNR of all users from all APs, denoted as a matrix S, current load on every [$N_u \times N_{AP}$], the dimensions of L_{AP} are [N_{AP}] in length US is [N_U] in width and height. As a result, the number of variables or features that are considered in the data point space. [N_u x N_{AP} + N_u x N_{AP}].

• APs are represented by the discrete actions in a finite set denoted by A. Thus, [N_U] is the correct value for the size of the discrete action space.

• The reward function, specified by r: S × A → R, is defined as ($r_{\mu,\,a}$ $K_{\mu,\,a}$)/ Rμ, in which $r_{\mu,\,a}$ denotes achievable data rate between the user and AP in accordance with Eqs. (6) and (10), R denotes the needed data rate related to the user, and $K_{\mu,a}$ denotes time slot allocation between the user and the AP in accordance with Eqs. (6) and (10).

$$k_{\mu,\alpha} = \frac{1}{\sum_a g_{\mu,\alpha}}, \quad (14)$$

• If user is connected to AP, then the corresponding binary variable will have a value of 1, and if they are not, then the value will be 0. Throughput and user satisfaction are taken into account while designing the incentive function. In particular, it is defined by the time slot allocation (denoted by $K_{\mu,\,a}$) between the AP and the user, the needed data rate of the user (denoted by), and the data rate of AP.

• A concession factor of 0.9 has been established P: S x A x S → R, represent the distribution of transition probabilities, and ρ_0: S → R represent the distribution of initial states s_π.

In this research, an MLP with parameters was used as policy network for AP assignment problem. Thus, we can write the policy as (at|st;θ). state-action (Q_π), value function (V_π), and advantage function (A):

$$Q_\pi(s_t, a_t) = E_{s_t+1, a_{t+1}, \ldots}[\sum_{l=0}^{\infty} \gamma^t r(s_{t+1}, a_{t+1})],$$
$$V_\pi(s_t) = E_{s_t+1, a_{t+1}, \ldots}[\sum_{l=0}^{\infty} \gamma^t r(s_{t+1}, a_{t+1})],$$
$$A_\pi(s, a) = Q_\pi(s, a) - V_\pi(s),$$
$$where$$
$$a_t \sim \pi(a_t|s_t;\theta), s_{t+1} \sim P(s_{t+1}|s_t, a_t)\ for\ t \geq 0. \quad (15)$$

Trust region policy optimization (TRPO) can be defined as a gradient descent-based technique for scalable policy optimization, and it is used throughout the training process to optimise the policy parameter of the MLP and maximise the predicted discounted return. Policy gradient methods are more stable throughout training and don't require a model, in contrast to value iteration approaches [19-21].

Furthermore, given certain conditions. TRPO guarantees improvements that grow steadily over time. Specifically, TRPO algorithms use gradient descent to directly learn policy (a|s; θ), while confining update of to a fixed amount at each step, indicating that every update to the policy during each iteration of TRPO results in a superior policy. To do this, TRPO limits the size of each iterative update by considering the KL divergence between the current and target policies. In TRPO, the optimization problem is stated as:

$$maximize\ E_{s\sim\rho\theta_{old}}, a \sim q\,[\frac{\pi_\theta(a|s)}{q_{(a|s)}}\, Q_{\theta\,old}\,(s,a)]$$
$$subject\ to$$
$$E_{s\sim\rho\theta_{old}}\,[D_{KL}(\pi_{\theta_{old}}\,(.|s)||\pi_\theta(.|s))] \leq \delta \quad (16)$$

This equation represents the optimization problem in TRPO, where δ is a parameter that can be adjusted.

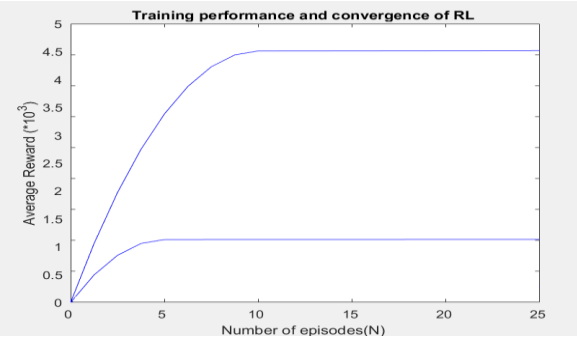

Figure. 4 Performance of RL with gamma values of 0.9 (red) and 0.7 (blue)

In order to measure how well RL performs and how far it has come, we can look at how much on average we have gained across numerous episodes, discounted at various rates (gamma). In this case, we are contrasting the efficacy of RL with gamma values of 0.9 (red) and 0.7 (blue).

The first algorithm (denoted *) π^* θ (at|st). Lays out the procedures that must be followed to arrive at the best possible policy. A stochastic policy that is called $\pi^*(a|s; \theta)$ is created when training is finished, and it assigns a distribution of probabilities between each one of the states and potential actions. This indicates that, given a certain situation, the policy gives a likelihood of choosing a specific course of action. As mentioned in algorithm 2, this policy could be applied in real-time applications for the prediction of proper AP assignments, calculating rewards, and determine subsequent state. RL-based systems often have a reduced computational complexity at run-time in comparison with optimization-based techniques. We developed an optimization problem with objective function that is identical to reward function of RL approach and reviewed alternative solutions that have been presented in section III-B for giving a fair comparison between the two methods. The analysis of RL algorithm's training and convergence performance will take place in the next section.

1) Training performance and RL convergence

The TRPO technique, which limits the policy to an acceptable area of trust to guarantee consistent learning, has been used in this research. Training results for 5 users have been presented to keep the presentation manageable, however, similar tendencies can be noticed for 10 users. The average reward earned during training with the RL algorithm is shown in Fig. 4 for a range of discount factors (γ). It appears that after a specific number of the episodes, mean reward that has been acquired by the algorithm converges for both values, but there is a large disparity between the average's rewards obtained for different values. To be more precise, when $\gamma = 0.9$ is used, the average reward is around three times as high as when $\gamma = 0.7$ is used. Because of this, although it will take longer for the method to converge, $\gamma = 0.9$ was selected for the simulation [22-24]. The convergence speed of the algorithm will be increased in future work by including the idea of knowledge transfer into the algorithm. Since TRPO is a model-free algorithm, it can learn the best policy without making any assumptions about the surrounding environment (transition and reward functions). Consequently, TRPO has a higher sample complexity but requires less hyperparameter adjustment. Exploration in TRPO depends on both the beginning conditions and the training process, as it is an on-policy approach that samples actions depending on the most recent version of its stochastic policy. When it comes to dealing with TRPO's complicated samples, many different approaches have been offered in the literature. However, the focus of this study is on RL-based load balancing for hybrid LiFi Wi-Fi networks, hence the simplest TRPO form was investigated. The algorithm's performance and rate of convergence will be improved in future work by incorporating knowledge transfer [25].

B. Other load balancing method

The aim of RL system that has been suggested in this work is consistent with the formulation of an optimization problem concerning AP assignment in hybrid LiFi Wi-Fi networks in the following way:

$$Max_{g_{\mu,\alpha}\, k_{\mu,\alpha}} = \sum_{\mu \in U} \sum_{\alpha \in AP} g_{\mu,\alpha} log\left(\frac{r_{\mu,\alpha} k_{\mu,\alpha}}{R_{\mu}}\right)$$
$$s.t. \sum_{\mu \in U} g_{\mu,\alpha} k_{\mu,\alpha} = 1 \; \forall \alpha \in AP \quad (17)$$

$$\sum_{\alpha \in AP} g_{\mu,\alpha} = 1 \; \forall \mu \in U$$
$$g_{\mu,\alpha} \in \{0,1\}, k_{\mu,\alpha} \in [0,1], \forall_{\mu} \in U, \forall_{\alpha} \in AP \quad (18)$$

R represents the needed data rate for the user, U denotes a set of users, AP represents a set of APs (which might contain Wi-Fi as well as LiFi APs), r denotes the achievable data rate between the user and AP, as specified by Eqs. (5) and (9), and AP and user have different achievable data rates. The optimization problem for the AP assignment in hybrid LiFi Wi-Fi networks is described in this formulation. If g, a = 1, then the user is connected to the AP; or else, they are not, according to the Boolean optimization variable g. The optimization variable defines allocating time slots to users who are all linked to AP. This optimization problem represents a MINLP (i.e. mixed integer nonlinear programming) problem with intractable mathematics, hence there isn't a closed-form solution available. By figuring out how to solve the MINLP problem that is specified by Eq. (13), there are three different approaches to optimize AP assignment in a hybrid LiFi Wi-Fi network.

Exhaustive search:

Through systematically listing all alternative actions for specified objective function and after that choosing the action that best satisfies the objective function, the approach obtains the optimal performance, yet at a high complexity cost. The room dimension in this scenario limits the number of users and APs, which keeps the computing complexity under control. Exhaustive searches use the goal function of Eq. (13) to determine a performance upper bound.

Iterative algorithm: Iterative algorithms could be used to tackle the optimization problem stated by Eq. (11). With the use of Lagrangian dual decomposition on the objective function, the optimal g value, a could be determined as follows:

$$g_{\mu,a} = \begin{cases} 1, \alpha = argmax\ (log\ (\frac{r_{\mu,\alpha}}{d_{\mu}}) \\ \qquad -\lambda_{\mu} - \omega_{\alpha}) \\ 0, otherwise \end{cases} \quad (19)$$

Following a similar procedure as described [25], the process of AP assignment and resource allocation has been carried out by calculating the Lagrangian multipliers $\lambda\mu$ and ω_a and using them to determine the optimal value of$g_{\mu,a}$.

RL-based approach: The suggested RL-based method immediately learns the optimal policy without assessing a model of the environment, which makes tuning the hyperparameters fairly simple. Through sampling actions following the most recent iteration of its stochastic policy, which is dependent on the initial conditions as well as training procedure, the RL algorithm explores the action space throughout training. The RL-based method offers a computationally effective solution while having a higher sample complexity.

Signal strength strategy (SSS): is a technique used to enhance the AP selection process in a network that combines LiFi and Wi-Fi. Due to differences between LiFi and Wi-Fi in terms of bandwidth and receiver noise, received signal intensity alone is not sufficient to determine the quality of the channel. Due to its importance, the signal-to-noise ratio (SNR) is the criterion relied upon by SSS. Assuming a set of access points (APs) consisting of one Wi-Fi AP and four LiFi APs, the SSS approach defines the objective function for a particular user as optimising the SNR between that user and the AP of interest, with the caveat that the AP of interest must be included in the set of APs.

$$\max_{\alpha} SNR_{\mu,\alpha} \quad s.t\ \alpha \in AP \quad (20)$$

SNR value between user μ and AP α, which are estimated sing Eqs. (9) & (4) for Wi-Fi and Li-Fi APs, respectively, is denoted by SNRμ,α.

4. Evaluation and discussion of performance

As shown in Fig. 1, we have selected a standard indoor space with dimensions of 5 x 5 x 3 meters, where a single Wi-Fi AP provides full coverage and four LiFi APs provide partial coverage. While our work can be scaled to accommodate more APs and users in a larger room, we chose a common situation to demonstrate the utility of RL in achieving an optimum AP selection policy for hybrid Li-Fi Wi-Fi networks and to offer an upper bound performance via exhaustive search. Here, it is assumed that all of the optical attocell utilise the same band-width, leading to interference that is merely dismissed as background noise in the overlapping regions. High data rates are anticipated for user access to HD videos, which is represented as a Poisson process at a rate of 50 Mbps. We have evaluated two mobility models, RWP and HRWP, with users spread out evenly over the room, as detailed in section 2. C. To keep things straightforward, we've assumed that TDMA is utilised to handle multiple connections and that users can only connect to a single access point (either the LiFi AP or the Wi-Fi AP).

Table 3. Parameters of system

System Parameter	Value
User Speed	1m/sec
Multiple access technology	TDMA
Requested data rate, R_{μ}	Poisson with 50Mbps
Policy	MLP, 2 layers of 64
Episode length, E	1,000
Gym environment	Li-Fi Wi-Fi network
Discount factor, γ	0.90
Maximum KL divergence, δ	0.010
Number of APs	4LiFi + 1Wi-Fi
RL Algorithm	TRPO
Dimensions of the Room	5x5x3 m^3
No. of Users, N_u	5,10
Wi-Fi AP location	(2.5m, 2.5m)
Li-Fi AP locations	(±1.25m,±1.25m)
User distribution	Uniform
Mobility Model	RWP, HRWP

The CU is equipped with an RL agent that determines load balancing decisions and can access both LiFi and Wi-Fi APs. In order to transmit the information of user states to the CU, it is assumed that a feedback link free of errors exists. The simulation was developed using Python 3.7, and Open AI Gym environment was created for the hybrid LiFi Wi-Fi network being studied [26]. The TRPO algorithm was implemented using the stable-baseline GitHub repository [27]. The reported results are an average higher than 200 episodes, and Table 3 summarizes system parameters that were considered for the simulations.

We compared the suggested RL approach to others, including exhaustive optimization, iterative

optimization, and SSS approaches, to determine how well it performed. Several criteria were used to make this comparison, including computational complexity, average network performance, fairness, user happiness, and the probability of capacity outages [28].

A. Performance matrix

The three performance measures are defined as follows: capacity outage probability, user satisfaction, and fairness.

- The satisfaction of a user, denoted by $S_{\mu,a}$, has been calculated as ratio of the achieved data rate for that user to required data rate, and can be represented as:

$$S_{\mu,a} = min\{1, \frac{k_{\mu,\alpha} r_{\mu,\alpha}}{R_{\mu}}\} \quad (21)$$

Here, R_{μ} denotes the demanded data rate by the user μ, modeled as Poisson process. The upper limit regarding user satisfaction has been 1, specifying that the user had accomplished their needed data rate.

- For fairness, Jain's fairness index [29] has been utilized.

Table 4. Computational complexity

Scheme	Complexity
SSS	$O(N_U N_{AP})$
Iterative	$O(N_U N_{AP} l)$
RL	$O(N_u^2 N_{AP} + N_U^2 + N_U N_{AP})$
Exhaustive	$O((N_{AP})^{Nu})$

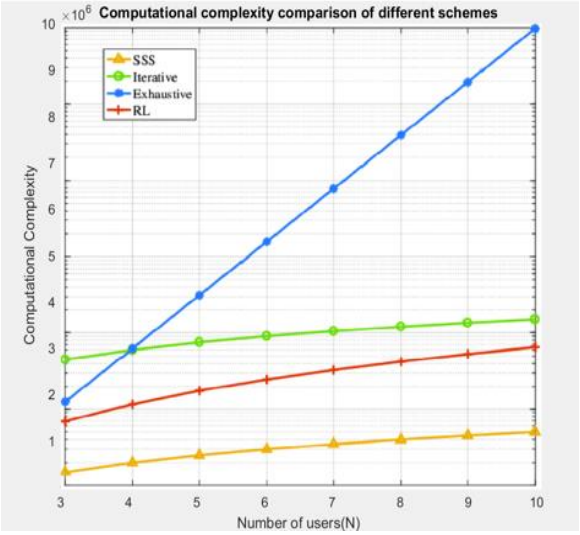

Figure. 5 Comparison of the computational complexity of various schemes

According to, it is possible to calculate the user fairness index.

$$\eta = \frac{(\sum_U S_{\mu,\alpha})^2}{N_U \sum_U (S_{\mu,\alpha})^2} \quad (22)$$

N_u represents the number of users

- The probability of capacity outage may be expressed by the following equation:

$$\Phi = Pr(k_{\mu,\alpha} r_{\mu,\alpha} < Ro) \quad (23)$$

The symbol "Ro" represents the average required throughput threshold value.

B. Complexity analysis

The Big-O notation is used here to quantify complexity because it establishes a ceiling for performance independent of hardware. The SSS method has a complexity of [30]. Because it selects AP with maximum SNR value out of all of the accessible APs (N_{AP}) Exhaustive search, the optimization method that relies on it, is computationally intensive because it investigates every conceivable pairing of users and APs, with a complexity of O ((N_{AP}) N_U). The complexity of an iteratively-solved optimization issue is estimated as O (N_{AP} N_U l), where I is the number of iterations needed to converge. The computational cost of the RL training phase is O (n2), where n is the number of states in MDP (n). To be clear, this is only the one-time investment required for training.

Table 5. Fairness

N_u	SSS	Iterative	Exhaustive	RL
5	116.23	174.04	239.90	218.81
10	73.12	98.72	138.41	110.69

Table 6. Average throughput of the network (Mbps)

N_u	SSS	Iterative	Exhaustive	RL
5	0.9998	0.9999	1.0000	1.000
10	0.9568	0.9743	0.9872	0.9727

The optimal strategy, here in the form of MLP, affects the complexity of the RL algorithm under real-time conditions. Part III lays out the parameters that govern the complexity of the system, including the size of the observation and action spaces. Because of this, RL has a complexity of O ($N_u^2 N_{AP} + N_U^2 + N_U N_{AP}$). The complexity of exhaustive optimization grows exponentially with the number of users, while complexity of SSS grows linearly. For RL, the difficulty scales quadratic ally with the number of users, while for the iterative approaches, the complexity scales linearly with I. Table 4 shows that RL has a significantly reduced complexity compared to exhaustive and iterative optimization. In Fig. 5 we illustrate the computational complexity of those approaches for varying numbers of users, considering that I = 30, as can be seen in the graph, RL has a lower complexity than both exhaustive and iterative approaches.

C. Number of users effect

Table 5 displays the typical throughput of a network with varying user loads. For a network with five users, the average throughput increases by 49.74% when using the iterative optimization-based AP assignment scheme instead of the standard solution set size (SSS). Moreover, RL and exhaustive optimization produce even more substantial improvements in average network throughput, with increases of 88.26% and 106.41% above SSS, respectively. For 10 users, both RL and exhaustive search outperform SSS by a significant margin, increasing average network throughput by 51.37 and 89.29%, respectively. Given that the average network throughput improvement over SSS for the iterative approach is only 34.24 percent, it is clear that RL is capable of providing a far larger improvement in throughput. It's important to remember that as the number of the users grows, network's performance saturates, limiting benefits that may be gained from using various load balancing techniques.

Table 6 displays the results of Eq. (17)'s calculations of Jain's fairness index for a range of user counts and load balancing strategies. When there are five users, the fairness value for all of the schemes is very near to 1. SSS, on the other hand, can only achieve fairness of 0.96 with 10 users, whereas RL and optimization both reach fairness of 0.97 in this scenario. Extensive optimization yielded the best possible fairness index of 0.97. Although the Jain's fairness index and the average network throughput present an overarching picture of the performance of the network, the CCDF of user happiness and the probability of capacity outage at a given throughput is supplied to shed light on how well these strategies are implemented.

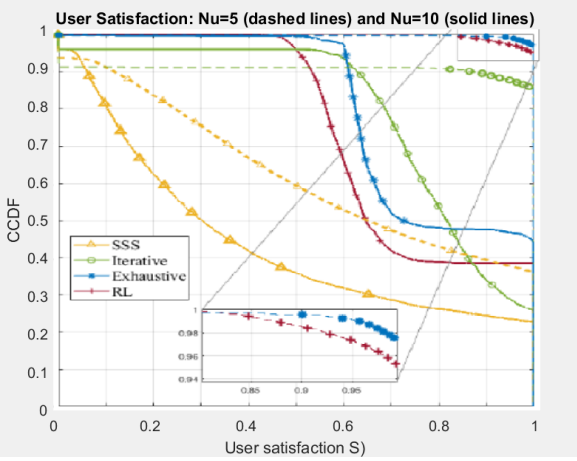

Figure. 6 User satisfaction: Nu=5 (dashed lines) and Nu=10 (solid lines)

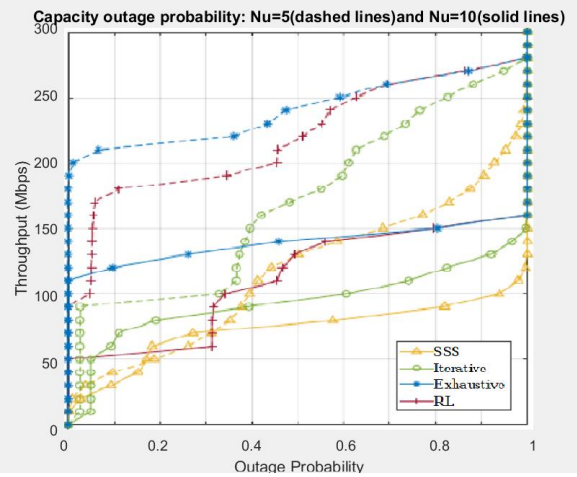

Figure. 7 Capacity outage probabilities: Nu=5 (dashed lines) and Nu=10 (solid lines)

Fig. 6 shows the CCDF (i.e. complementary cumulative distribution function) of user satisfaction for varying load balancing schemes and user counts. The happiness index is determined by solving for Eq. 16. With less than 40% and 25% of users being entirely satisfied for 5 and 10 users, respectively, the data show that the SSS-based system gives the worst customer satisfaction. For 5 users, the satisfaction percentage for the iterative optimization strategy is over 85%, but with 10 users, it declines dramatically to around 30%. Whereas, roughly 98% and 96% of 5 users are completely satisfied with the exhaustive optimization and RL techniques, respectively. With ten users, however, those percentages drop to about 45 and 40, respectively. Consequently, it can be stated that RL delivers close enough speed, while the exhaustive search approach provides the largest percentage of full user satisfaction. Also, it is seen that as the number of users grows, level of satisfaction decreases across the board.

Table 7. Avg. network throughput (Mbps)

Mobility	SSS	Iterative	Exhaustive	RL
RWP	119.230	184.040	250.90	220.81
HRWP	129.230	200.280	297.86	210.85

Table 8. Fairness

Mobility	SSS	Iterative	Exhaustive	RL
RWP	0.99980	0.99990	1.000	1.000
HRWP	0.99900	0.99680	1.000	1.000

Depending on the load balancing scheme and the number of users, the probability of a capacity outage for a throughput threshold of Ro is shown in Fig. 7. For 5 users, the average network throughput for exhaustive search and RL can reach up to 210 Mbps and 180 Mbps, respectively, when the system permits 10% of users to be in capacity outage. SSS shows the lowest throughput value of roughly 40 Mbps, while the iterative method performs poorly with a throughput value of less than 100 Mbps. Whenever the chance of a capacity outage rises above 0.6, the benefits of using RL or thorough optimization are nearly the same. In addition, the throughput value drops drastically for a given capacity outage probability when the number of the users rises from 5 to 10. Exhaustive search delivers the best potential throughput for a given probability of capacity outage, with RL coming in a close second. Lastly, when the number of the users grows, performance of all of the schemes degrades, while RL's performance is quite near to that of exhaustive search.

D. Mobility model effect

We also considered how specific user activities, such as the four attraction points around the 4 LEDs in the HRWP mobility model, would have an impact

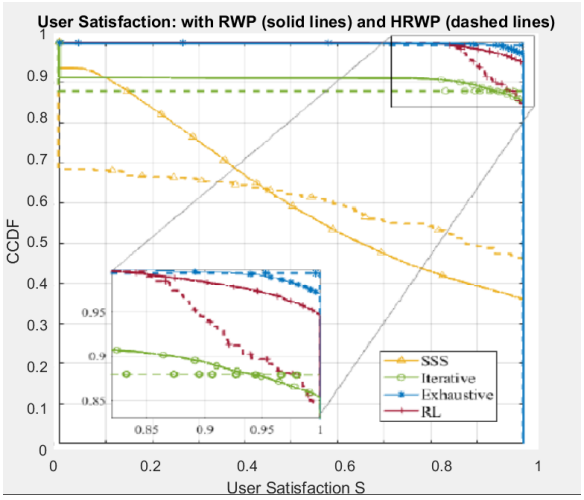

Figure. 8 User satisfaction: with RWP (solid lines) and HRWP (dashed lines)

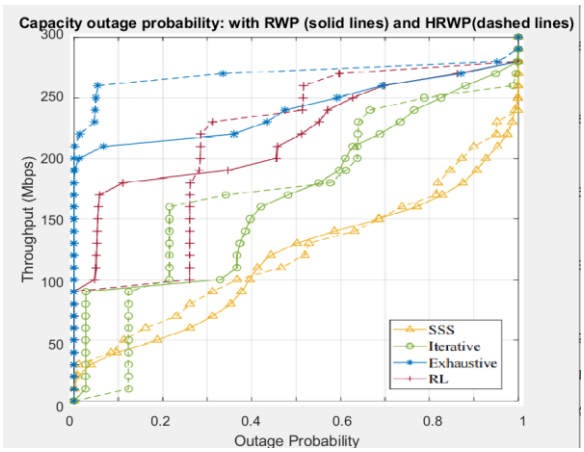

Figure. 9 Capacity outage probability: with RWP (solid lines) and HRWP (dashed lines)

on the outcomes throughout this analysis. Table7 displays the typical network throughput for five users using 2 mobility models. The two mobility models displayed the same overall performance trends, with exhaustive optimization, succeeded by RL and iterative optimization, increasing average network throughput the greatest in comparison to the SSS method. Jain's fairness index for five users across various mobility scenarios is shown in Table 8. It is evident that all plans for both forms of mobility achieved complete fairness ($\eta = 1$) for a group of 5 users.

In addition, this research examined how different mobility models affected the overall system efficiency. Both the high-resolution wheeled platform (HRWP) and the reduced-complexity wheeled platform (RWP) mobility models were investigated. Table 7 provides a high-level overview of the average network throughput for 5 users across both models, revealing a consistent performance profile. The average network throughput was

improved the most from the SSS method by using the exhaustive optimization technique, followed by RL and iterative optimization. In Table 8, we can see that all five schemes achieved full fairness (η =1) for five users under the two models of mobility.

Fig. 8 depicts the cumulative distribution function (CCDF) regarding user satisfaction for 5 users under both models of mobility, revealing that user satisfaction varies considerably between the two. Under HRWP, roughly 45% of SSS users may obtain 100% satisfaction, while under RWP, that number drops to 35%. Users are somewhat more satisfied with HRWP after utilising exhaustive search. Additionally, RL provides HRWP performance that is on par with the iterative optimization method. Iterative optimization method just ensures UXSatisfaction for 90% of users, whereas RL guarantees 0.85 UXSatisfaction for all users across both mobility models. This indicates that with the iterative method, at least 10% of users never achieve the desired UXSatisfaction. Furthermore, in the HRWP paradigm, SSS can only guarantee user happiness for 70% of users.

Fig. 9 shows the percentage of times a specific throughput Ro will suffer a capacity issue for five users using two distinct mobility models. It demonstrates that with exhaustive search, the throughput that can be achieved with a given capacity outage probability improves in the case of HRWP. The functionality of SSS is nearly identical between RWP and HRWP. The iterative algorithm and RL act comparably for the 2 models of mobility, with higher throughput being supported for smaller capacity outage probability values in RWP models and for larger outage probability values in HRWP models. The two mobility models' performance trends for the four load balancing techniques are stable, and RL provides resilience in the face of a variety of user activities.

E. Discussion and comparison

The outcomes of a hybrid system built on reinforcement learning (RL) and standard solution set size (SSS) access technology known as TDMA, as achieved through MATLAB code. Five users' throughput and user satisfaction scores were compared to the system. The RL-based hybrid system achieved an SSS of 180 Mbps and a throughput of up to 210 Mbps, according to the data. This shows that the system was able to deliver high data transfer rates while making effective use of the resources at hand. Furthermore, a 100% user satisfaction rate was obtained, indicating that the system either matched or beyond users' expectations in terms of functionality and level of service. Overall, the findings highlight the effectiveness of the RL-based hybrid system with TDMA access technology in achieving high throughput, efficient resource allocation, and user satisfaction in a multi-user scenario. Table 9 shows the comparison between the performance of the proposed hybrid system and the performance of the hybrid system in previous works.

Table 9. Comparison between the performance of the proposed hybrid system and the performance of the hybrid system in previous works

Ref	[17]	[28]	**Proposed System**
Year	2019	2020	2023
Algorithm	Fuzzy Logic	RL Method	Reinforcement Learning (RL) standard solution set size (SSS)
Access technology	TDMA	FDMA	TDMA
System method	one Wi-Fi AP and a four Li-Fi APs.	Hybrid Li-Fi/Wi-Fi network with four Li-Fi APs and one Wi-Fi AP	Hybrid system Based on Reinforcement Learning
Throughput of 5 users	95Mbps	110.69 Mbps 175 Mbps	RL can reach up to 210 Mbps and 180 Mbps SSS
User satisfaction	……..	70%	100%

5. Conclusion

For hybrid LiFi-Wi-Fi networks, this article suggests a dynamic load balancing strategy based on RL. Throughput should be maximised on a long-term basis on a system-wide average, with QoS guarantees being a secondary concern. The TRPO method is used to train the RL to learn the best course of action for AP allocation in every given scenario. Simulations have demonstrated that the suggested approach is more effective than both approaches the standard SSS scheme and iterative algorithm. RL algorithm surpasses SSS and iterative methods in terms of throughput for 5 users in the RWP mobility model by 87.36% and 24.27%, respectively. The exhaustive optimization approach achieves the highest performance, but it is computationally intensive and therefore not applicable to most real-

world situations. The RL scheme is as effective as the exhaustive search, but simpler, in terms of user happiness. Furthermore, the suggested RL scheme's performance tendencies are consistent across a wide range of mobility models, demonstrating its robustness. For improving the algorithm's rate of convergence and efficiency, it will be necessary to incorporate knowledge transfer into future projects.

Conflicts of interest

The authors declare that they have no conflicts of interest. This research was conducted independently and without funding from any external sources.

Author contributions

Conceptualization, Zeena Mustafa; methodology, Zeena Mustafa ; software, Ekhlas Kadhum Hamza; validation, Zeena Mustafa ; formal analysis, Zeena Mustafa ; investigation, Zeena Mustafa; resources, Zeena Mustafa; data curation, Zeena Mustafa; writing—original draft preparation, Zeena Mustafa; writing—review and editing, Zeena Mustafa ; visualization, Zeena Mustafa ; supervision, Ekhlas Kadhum Hamza; project administration, Ekhlas Kadhum Hamza; funding acquisition, Not funded.

Acknowledgments

The authors acknowledge The Department of Control and Systems Engineering at the University of Technology in Iraq.

List of notations

Abbreviation	Description
HLOS	Channel gain of the line-of-sight (LOS) component.
m	Path loss exponent.
A_{PD}	Area of the photodetector.
d	Distance between the transmitter and receiver.
φ	Angle of arrival of the LOS component.
g_f	Gain factor associated with the frequency response.
g_c	Gain factor associated with the coverage area.
ψ	Angle of departure of the LOS component.
Ro	the average required throughput threshold value.
NLOS	non-line-of-sight
R	data rate for the user
TRPO	Trust Region Policy Optimization
SNR	Signal-to-Noise Ratio
$K_{\mu,\alpha}$	denotes time slot allocation between the user and the AP
LFS	the free space loss
fc	cut-off frequency
PSD	power spectral density
S(μ,a)	satisfaction of a user
n	Refractive index of the medium.
H_{NLOS}	Channel gain of the non-line-of-sight (NLOS) component
ρ	Reflectance of the surrounding environment.
A_PD	Area of the photodetector.
f	Frequency of the signal
ΔT	Delay spread
A_room	Area of the room or space
fc	Carrier frequency.
G(μ,α)(f)	The channel gain for WiFi
μ	Path loss exponent
α	Shadowing factor
L(d)	Path loss as a function of distance d
$H_{LiFi(\mu,\alpha)}$	The complete optical channel for Li-Fi, consisting of the LOS (HLOS) and NLOS (HNLOS) components.
P_{opt}	Optical power transmitted by the Li-Fi access point.
N_{LiFi}	Total noise power in the Li-Fi system.
B_{LiFi}	Bandwidth of the Li-Fi system.
$f(x_d, y_d)$	Distribution function for destination points (xd, yd).
ξ	Intensity parameter.
A_a	Area of the attraction point.
x_m	Half the length of the attraction point area in the x-direction.
y_m	Half the length of the attraction point area in the y-direction.
$x_{a,min}$	Minimum x-coordinate of the attraction point area.
$y_{a,min}$	Minimum y-coordinate of the attraction point area.

References

[1] S. Aboagye, T. M. N. Ngatched, O. A. Dobre, and A. Ibrahim, "Joint Access Point Assignment and Power Allocation in Multi-Tier Hybrid RF/VLC HetNets", *IEEE Trans. Wirel. Commun.*, Vol. 20, No. 10, pp. 6329–6342, 2021.

[2] M. Rahaim, I. Abdalla, M. Ayyash, H. Elgala, A. Khreishah, and T. D. C. Little, "Welcome to the CROWD: Design Decisions for Coexisting Radio and Optical Wireless Deployments", *IEEE Netw.*, Vol. 33, No. 5, pp. 174–182, 2019.

[3] Q. Ling, L. B. L. Baoshan, and D. Yongxing, "Progress report on visible light communication in intelligent transportation environment", *J.*

Phys. Conf. Ser., Vol. 1168, No. 2, 2019.

[4] Y. S. Hussein and A. C. Annan, "Li-Fi Technology: High data transmission securely", *J. Phys. Conf. Ser.*, Vol. 1228, No. 1, 2019.

[5] A. R. Ndjiongue, T. M. N. Ngatched, O. A. Dobre, and A. G. Armada, "VLC-based networking: Feasibility and challenges", *IEEE Netw.*, Vol. 34, No. 4, pp. 158–165, 2020.

[6] A. L. Oros, "5. The Next Frontier", *Norm. Japan*, pp. 122–148, 2020.

[7] X. Wu, M. D. Soltani, L. Zhou, M. Safari, and H. Haas, "Hybrid LiFi and WiFi Networks: A Survey", *IEEE Commun. Surv. Tutorials*, Vol. 23, No. 2, pp. 1398–1420, 2021.

[8] C. Sun, J. Wang, X. Gao, and Z. Ding, "Networked Optical Massive MIMO Communications", *IEEE Trans. Wirel. Commun.*, Vol. 19, No. 8, pp. 5575–5588, 2020.

[9] G. Mirzaeva, G. C. Goodwin, B. P. McGrath, C. Teixeira, and M. E. Rivera, "A Generalized MPC Framework for the Design and Comparison of VSI Current Controllers", *IEEE Trans. Ind. Electron.*, Vol. 63, No. 9, pp. 5816–5826, 2016.

[10] X. Wu, M. Safari, and H. Haas, "Access Point Selection for Hybrid", In: *Proc. of IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Vol. 65, No. 12, pp. 5375–5385, 2017.

[11] S. Nafea and E. K. Hamza, "Path loss Optimization in WIMAX Network using Genetic Algorithm", *Iraqi J. Comput. Commun. Control Syst. Eng.*, No. October, pp. 24–30, 2020.

[12] Y. Wang and H. Haas, "A comparison of load balancing techniques for hybrid LiFi/RF networks", In: *Proc. of VLCS 2017 - Proc. 4th ACM Work. Visible Light Commun. Syst. co-located with MobiCom 2017*, pp. 43–47, 2017.

[13] T. A. Ali and A. M. Hasan, "Low-Cost MEMS-Based NARX Model for GPS- Denied Areas", *The Iraqi Journal of Computers, Communications, Control, and Systems Engineering*, No. 4, pp. 58–70, 2020.

[14] Y. Wang, D. A. Basnayaka, X. Wu, and H. Haas, "Optimization of Load Balancing in Hybrid LiFi/RF Networks", *IEEE Trans. Commun.*, Vol. 65, No. 4, pp. 1708–1720, 2017.

[15] X. Wu and H. Haas, "Access point assignment in hybrid LiFi and WiFi networks in consideration of LiFi channel blockage", *IEEE Work. Signal Process. Adv. Wirel. Commun. SPAWC*, Vol. 2017-July, No. 978, pp. 1–5, 2017.

[16] W. Ma and L. Zhang, "QoE-Driven optimized load balancing design for hybrid LiFi and WiFi Networks", *IEEE Commun. Lett.*, Vol. 22, No. 11, pp. 2354–2357, 2018.

[17] M. A. Hussein and E. K. Hamza, "Secure Mechanism Applied to Big Data for IIoT by Using Security Event and Information Management System (SIEM)", *Int. J. Intell. Eng. Syst.*, Vol. 15, No. 6, pp. 667–681, 2022, doi: 10.22266/ijies2022.1231.59.

[18] X. Wu and H. Haas, "Load Balancing for Hybrid LiFi and WiFi Networks: To Tackle User Mobility and Light-Path Blockage", *IEEE Trans. Commun.*, Vol. 68, No. 3, pp. 1675–1683, 2020.

[19] S. Sharmin, I. Ahmedy, and R. M. Noor, "An Energy-Efficient Data Aggregation Clustering Algorithm for Wireless Sensor Networks Using Hybrid PSO", *Energies*, Vol. 16, No. 5, 2023.

[20] R. Shanmughasundaram, S. P. Vadanan, and V. Dharmarajan, "Li-Fi Based Automatic Traffic Signal Control for Emergency Vehicles", In: *Proc. of 2018 2nd Int. Conf. Adv. Electron. Comput. Commun. ICAECC 2018*, pp. 1–5, 2018.

[21] T. H. Nasser, E. K. Hamza, and A. M. Hasan, "MOCAB / HEFT Algorithm of Multi Radio Wireless Communication Improved Achievement Assessment", *Bulletin of Electrical Engineering and Informatics*, Vol. 12, No. 1, pp. 1–8, 2023.

[22] S. Aboagye, A. Ibrahim, T. M. N. Ngatched, A. R. Ndjiongue, and O. A. Dobre, "Design of Energy Efficient Hybrid VLC/RF/PLC Communication System for Indoor Networks", *IEEE Wirel. Commun. Lett.*, Vol. 9, No. 2, pp. 143–147, 2020.

[23] L. Zhou, A. Swain, and A. Ukil, "Reinforcement Learning Controllers for Enhancement of Low Voltage Ride through Capability in Hybrid Power Systems", *IEEE Trans. Ind. Informatics*, Vol. 16, No. 8, pp. 5023–5031, 2020.

[24] M. Kamel, R. Dai, Y. Wang, F. Li, and G. Liu, "Data-driven and model-based hybrid reinforcement learning to reduce stress on power systems branches", *CSEE J. Power Energy Syst.*, Vol. 7, No. 3, pp. 433–442, 2021.

[25] S. Aziez, E. Hamza, F. Hummadi, and A. Sabry, "Implementation of Radar Cross-Sections Model for Targets With Different Scattering Centers", *Eastern-European J. Enterp. Technol.*, Vol. 5, No. 9–119, pp. 54–60, 2022.

[26] A. Kumar, S. Sarkar, and C. Pradhan, "Malaria Disease Detection Using CNN Technique with SGD, RMSprop and ADAM Optimizers", *Stud. Big Data*, Vol. 68, No. April, pp. 211–230, 2020.

[27] B. Jagadeeswari, C. S. Anusha, and D. M. M. Preethi, "Audio Transmission using Li-Fi Technology", *Int. J. Trend Sci. Res. Dev.*, Vol. 3,

No. Issue-3, pp. 1008–1011, 2019.
[28] R. Ahmad, M. D. Soltani, M. Safari, A. Srivastava, and A. Das, "Reinforcement Learning Based Load Balancing for Hybrid LiFi WiFi Networks", *IEEE Access*, Vol. 8, pp. 132273–132284, 2020.
[29] D. Wu, B. Liu, Q. Yang, and R. Wang, "Social-aware cooperative caching mechanism in mobile social networks", *J. Netw. Comput. Appl.*, Vol. 149, No. July 2019, p. 102457, 2020.
[30] F. Paquin, J. Rivnay, A. Salleo, N. Stingelin, and C. Silva, "Multi-phase semicrystalline microstructures drive exciton dissociation in neat plastic semiconductors", *J. Mater. Chem. C*, Vol. 3, No. 3, pp. 10715–10722, 2015.